

University of Groningen

The hierarchical assembly of galaxies and black holes in the first billion years: Predictions for the era of gravitational wave astronomy

Dayal, Pratika; Rossi, Elena M.; Shiralilou, Banafsheh; Piana, Olmo; Choudhury, Tirthankar Roy; Volonteri, Marta

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
[10.1093/mnras/stz897](https://doi.org/10.1093/mnras/stz897)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Dayal, P., Rossi, E. M., Shiralilou, B., Piana, O., Choudhury, T. R., & Volonteri, M. (2019). The hierarchical assembly of galaxies and black holes in the first billion years: Predictions for the era of gravitational wave astronomy. *Monthly Notices of the Royal Astronomical Society*, 486(2), 2336-2350.
<https://doi.org/10.1093/mnras/stz897>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

The hierarchical assembly of galaxies and black holes in the first billion years: predictions for the era of gravitational wave astronomy

Pratika Dayal¹,¹★ Elena M. Rossi,² Banafsheh Shiralilou,² Olmo Piana,¹ Tirthankar Roy Choudhury³ and Marta Volonteri⁴

¹Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, the Netherlands

²Leiden University, Oort Building, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands

³National Centre for Radio Astrophysics, Tata Institute of Fundamental Research, Pune 411007, Maharashtra, India

⁴Institut d'Astrophysique de Paris, Sorbonne Universités, UPMC Univ Paris 6 et CNRS, UMR 7095, 98 bis bd Arago, F-75014 Paris, France

Accepted 2019 March 22. Received 2019 March 12; in original form 2018 October 24

ABSTRACT

In this work, we include black hole (BH) seeding, growth, and feedback into our semi-analytic galaxy formation model, *Delphi*. Our model now fully tracks the accretion- and merger-driven hierarchical assembly of the dark matter halo, gas, stellar, and BH masses of high-redshift ($z \gtrsim 5$) galaxies. We explore a number of physical scenarios that include (i) two types of BH seeds (stellar and those from direct collapse BH); (ii) the impact of reionization; and (iii) the impact of instantaneous versus delayed galaxy mergers on the baryonic growth. Using a minimal set of mass- and z -independent free parameters associated with star formation and BH growth, and their associated feedback, and including the suppressed BH growth in lower mass galaxies, we show that our model successfully reproduces all available data sets for early galaxies and quasars. While both reionization and delayed galaxy mergers have no sensible impact on the evolving ultraviolet luminosity function, the impact of the former dominates in determining the stellar mass density for observed galaxies as well as the BH mass function. We then use this model to predict the *LISA* detectability of merger events at high- z . As expected, mergers of stellar BHs dominate the merger rates for all scenarios and our model predicts an expected upper limit of about 20 mergers using instantaneous merging and no reionization feedback over the 4-yr mission duration. Including the impact of delayed mergers and reionization feedback provides about 12 events over the same observational time-scale.

Key words: Galaxies: high-redshift – formation – evolution – star formation – quasars: super massive black holes; gravitational waves.

1 INTRODUCTION

The detection of gravitational waves (GWs), from merging stellar black hole (BH) binaries and recently from the mergers of binary neutron stars, has opened a new observable window on to the low- z Universe. Over the next decade, we expect the *Laser Interferometer Space Antenna* (*LISA*) to detect GW signals from merging massive (10^4 – $10^6 M_\odot$) BH binaries at the centre of galaxies from redshifts as high as $z \sim 20$, if such BHs are already in place at those early epochs. These observations perfectly complement galaxy surveys, using the *Hubble* and Subaru telescopes and the forthcoming *James Webb Space Telescope* (*JWST*) and the European Extremely Large telescope, which aim at yielding tantalizing glimpses of the formation of the earliest galaxies in the era of cosmic dawn (for a recent review, see Dayal & Ferrara 2018).

Predictions for such observations naturally require theoretical models that can consistently and simultaneously reproduce existing galaxy and BH observations before predictions are made for higher redshifts. In this work, we introduce BH seeding, growth, and feedback into the *Delphi* semi-analytic model of galaxy formation. This model now fully tracks the hierarchical assembly of the dark matter, baryonic and BH components of early galaxies including the impact of feedback associated with star formation, BHs, and reionization. We use the results of our model to forecast the BH merger rate and the associated GW signature for *LISA* between $z \sim 5$ and 20. This redshift window covers the formation epoch of supermassive black hole (SMBH) seeds, whose nature and properties are currently outstanding questions. Three main formation channels are currently discussed (see also Section 2.1) that involve: (i) the first generation of massive metal-free Population III stars (PopIII; e.g. Madau & Rees 2001); (ii) the monolithic collapse of gas in assembling protogalaxy (e.g. Loeb & Rasio 1994); and (iii) the core collapse of the first ultradense nuclear star clusters

* E-mail: p.dayal@rug.nl

(e.g. Devecchi & Volonteri 2009). These channels differ in terms of the expected seed mass and ‘birth’ rate with PopIII remnants being the lightest ($\sim 100 M_\odot$) and most frequent and BHs resulting from gas collapse (direct collapse black holes; DCBHs) being the most massive and rarest (see Latif & Ferrara 2016; Hartwig 2018, for recent reviews). Therefore, catching the GW signals from merging BHs between $z \sim 5$ and 20 can shed unique insights on SMBH infancy (Colpi 2018). Electromagnetic searches of SMBH seeds probe the fraction of the SMBH population in an active phase and provide their luminosity, therefore shedding light on the combination of their masses and accretion rates. An alternative but still indirect measure of the SMBH mass, which is one of quantities needed to test seed formation scenarios, can be estimated with spectroscopic observations. On the other hand, *LISA* can detect and directly provide the masses and spins for quiescent SMBHs, as long as they are in a coalescing binary. Therefore, these two approaches are truly complementary and both are necessary in order to obtain a full view of the SMBH seed population in the early Universe.

In this paper, we focus on BH–BH mergers that ensue after the merger of two galaxies, both hosting a central BH, rather than the merger of SMBHs that are born in a binary (Hartwig, Agarwal & Regan 2018). While such calculations have been carried out by previous works (e.g. Sesana, Volonteri & Haardt 2007; Sesana et al. 2011; Barausse 2012), the strength of our model lies, both, in the minimal set of free parameters used for star formation and BHs (and their associated feedback) as well as the number of physical scenarios explored that include (i) two types of BH seeds (stellar and those from DCBH); (ii) the impact of reionization; and (iii) the impact of instantaneous versus delayed galaxy mergers on the baryonic growth of galaxies. Our model matches the key observables both for high- z Lyman break galaxies (LBGs) – including the ultraviolet luminosity function (UVLF), the stellar mass function, the stellar mass density (SMD), stellar mass–halo mass relations, and the mass-to-light (M/L) ratios – and BHs – including their UVLF, the black hole mass function (BHMF), and the BH mass–stellar mass relations. The GW event rates predicted by this work are therefore benchmarked against all available high- z galaxy and BH data.

The cosmological parameters used in this work correspond to $(\Omega_m, \Omega_\Lambda, \Omega_b, h, n_s, \sigma_8) = (0.3089, 0.6911, 0.049, 0.67, 0.96, 0.81)$, consistent with the latest results from the *Planck* collaboration (Planck Collaboration 2015). We quote all quantities in comoving units unless stated otherwise and express all magnitudes in the standard AB system (Oke & Gunn 1983).

The paper is organized as follows. In Section 2, we detail our code for the galaxy–BH (co)-evolution, which we test against observations in Section 3. In Section 4, we simulate *LISA*’s performance in detecting our mock population of merging BHs and our results are summarized and discussed in Section 5.

2 THE THEORETICAL MODEL

This work is based on using the code *Delphi* (dark matter and the emergence of galaxies in the epoch of reionization), introduced in previous papers including Dayal et al. (2014), Dayal, Mesinger & Pacucci (2015), and Dayal et al. (2017a,b). In brief, *Delphi* uses a binary merger tree approach to jointly track the build-up of dark matter haloes, their baryonic components (both gas and stellar mass), and their star formation-driven spectra through cosmic time. We start by building merger trees for 550 $z = 4$ galaxies, uniformly distributed in the halo mass range of $\log(M_h/M_\odot) = 8\text{--}13.5$, up to $z = 20$. Each $z = 4$ halo is assigned a comoving number density by

matching to the dn/dM_h value of the $z = 4$ Sheth–Tormen halo mass function (HMF), and every progenitor halo is assigned the number density of its $z = 4$ parent halo; we have confirmed that the resulting HMFs are compatible with the Sheth–Tormen HMF at all z . In terms of feedback, so far, this model has focused on modelling the impact of (Type II) supernovae (SNII) and reionization feedback on the formation of early galaxies. In this work, we extend our model to include the impact of BH seeding, growth, and feedback on early galaxy formation. The very first progenitors of every $z = 4$ halo, which mark the start of its assembly (the so-called ‘starting leaves’), are assigned an initial gas mass that scales with the halo mass according to the cosmological ratio such that $M_g = (\Omega_b/\Omega_m)M_h$. Depending on their mass and redshift, such starting leaves can also be seeded with a BH as explained in Section 2.1. We start by calculating the star formation efficiency of a halo and the gas mass remaining after SN feedback (Section 2.2). If a halo hosts a BH, a part of the gas left after star formation and SN feedback can be accreted on to the BH and the impact of BH feedback is included as detailed in Section 2.3. At each step, we include both the impact of smooth accretion and mergers in assembling the halo, baryonic (gas and stellar mass), and BH mass as detailed in Sections 2.4 and 2.5. In our endeavour to build a model with minimal free parameters, we limit ourselves to two and four *mass- and z -independent* free parameters related to star formation and BHs, respectively.

2.1 Seeding haloes with BHs

We explore the two formation channels that yield the lightest (Pop III) and the most massive (DCBH) BH seeds. At variance with previous models for GWs from BH seeds that included only one type of BH seeds (except for the post-processed ‘mixed models’ in Sesana et al. 2011), here we consider the possibility of more than one BH formation mechanism operating in the (early) Universe, as generally expected (e.g. Valiante et al. 2016). These BH seeds are planted in the *starting leaves of any halo* as now detailed:

(i) *Heavy seeds*: First postulated as massive ($10^{3\text{--}5} M_\odot$) BH seeds to explain the presence of SMBHs at early cosmic epochs (e.g. Loeb & Rasio 1994; Bromm & Loeb 2003), DCBH formation models have been continually refined and developed over the past years (e.g. Begelman, Volonteri & Rees 2006; Begelman, Rossi & Armitage 2008; Regan & Haehnelt 2009; Shang, Bryan & Haiman 2010; Johnson et al. 2012; Latif et al. 2013; Agarwal et al. 2014; Dijkstra, Ferrara & Mesinger 2014; Ferrara et al. 2014; Habouzit et al. 2016). The current understanding from these works requires the following conditions to be met for a DCBH host: (i) the halo should have reached the atomic cooling threshold, with a virial temperature $T_{\text{vir}} \gtrsim 10^4$ K, for the gas to be able to cool isothermally; (ii) the halo should be metal-free to prevent gas fragmentation; and (iii) the halo should be exposed to a high enough ‘critical’ Lyman–Werner (LW) background ($J_{\text{crit}} = \alpha J_{21}$). Here, $\alpha > 1$ is a free parameter and J_{21} is the LW background expressed in units of $10^{-21} \text{ erg s}^{-1} \text{ Hz}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$ (see e.g. Sugimura, Omukai & Inoue 2014). Interested readers are referred to Dayal et al. (2017b) for complete details on how DCBHs are seeded in high- z haloes. In brief, we start by making the reasonable assumption that the starting leaves of any halo are metal-free by virtue of never having accreted metal-enriched gas. Further, we use the stellar population synthesis code STARBURST99 (Leitherer et al. 1999) to calculate the LW (11.2–13.6 eV) luminosity of each galaxy based on its entire star formation history. This is used to calculate the mean LW emissivity at a given redshift, $\epsilon_{\text{LW}}(z)$, by integrating over all galaxies present at that z .

Accounting for fluctuations in the background, most likely around galaxies and from the biased (i.e. clustered) distribution of galaxies, we identify the probability of the starting leaves being irradiated by LW intensity above a critical threshold value as detailed in Dayal et al. (2017b); for the calculations in this work, we explore values that range over an order of magnitude such that $\alpha = 30$ and 300.

The mass distribution of the seeds is uncertain and depends on the specific physical conditions at birth and on whether the intermediate state of a supermassive star is followed by a brief period of super-Eddington accretion on to the newly born BH (i.e. a quasi-star phase). For example, the supermassive star mass depends on the strength of the LW radiation that illuminates the birth site (Agarwal 2018; Latif 2018). On the other hand, the existence of a quasi-star phase and its outcome in terms of BH seed mass depend on internal rotation and on mass-loss in winds (Dotan, Rossi & Shaviv 2011; Fiacconi & Rossi 2016, 2017). We are therefore left with an uncertain SMBH seed mass that can be bracketed by 10^3 – $10^5 M_\odot$ in haloes below $10^9 M_\odot$. Given the halo masses and LW radiation thresholds used in this paper (cf. fig. 5 in Latif 2018), we randomly populate haloes in the top half of the calculated probability range with a DCBH seed of mass ranging between $10^{3-4} M_\odot$ (*‘light DCBH seeds’*): the number of haloes populated with such DCBHs is calculated by matching this DCBH mass function to the probabilistic one (obtained by multiplying the mass function of DCBH hosts with the hosting probability). In order to check the dependence of our results on the DCBH seeds mass used, we also show the *LISA* event rates expected for seed masses higher by an order of magnitude, ranging between $10^{4-5} M_\odot$. The results of this *‘heavy DCBH seeds’* model are shown in Section 4.3.

(ii) *Light seeds*: Stellar BH seeds of mass $\sim 10^2 M_\odot$ can be created by the collapse of PopIII stars in minihaloes with $M_h \sim 10^5 M_\odot$. Making the reasonable assumption that haloes collapsing from high ($\gtrsim 3.5$)- σ fluctuations in the primordial density field are most likely to host such seeds (e.g. Volonteri, Haardt & Madau 2003; Barausse 2012) results in the host haloes being more massive than $M_h \gtrsim 10^{7.2} M_\odot$ at $z \gtrsim 13$. However, given our halo mass range of $10^{8-13.5} M_\odot$, starting leaves have to be assigned these seeds by hand. In this work, we start by populating the starting leaves of any halo, which fulfil the DCBH criterion detailed above, with seed DCBHs. The starting leaves of haloes at $z \gtrsim 13$ that fulfil the light seed criterion, but do not contain a DCBH, are then populated with stellar BH seeds mass $M_{bh} = 150 M_\odot$.

The initial seed distribution obtained with this formalism is shown in Fig. 1. The cumulative number density of stellar BH seeds has a value of about $10^{-3.8} \text{ Mpc}^{-3}$ by $z \sim 4$. This clearly dominates over the cumulative number density of DCBH seeds that have a value of about $10^{-5.8}$ ($10^{-7.6}$) Mpc^{-3} by $z \sim 4$ for $\alpha = 30$ (300). Note that the models for ‘light seeds’ described above were inspired by early calculations that suggested that only one, very massive, PopIII star would form in a given halo, right at the centre of the potential well (e.g. Abel, Bryan & Norman 2002). More recent models favour a larger amount of fragmentation, leading to lighter and scattered PopIII remnants (Jeon et al. 2014; Smith et al. 2018) that are less suitable as SMBH seeds given the difficulty of both accreting material from their surroundings and finding the dynamical centre of the galaxy/halo via dynamical friction.

Therefore, we also run test cases with only DCBHs as a lower limit to the presence of SMBH seeds in primeval galaxies. In terms of GW signatures, the results are similar to the reference case but selecting only mergers between heavy seeds. In terms of general BH population, models including only DCBHs with $\alpha = 300$ fail

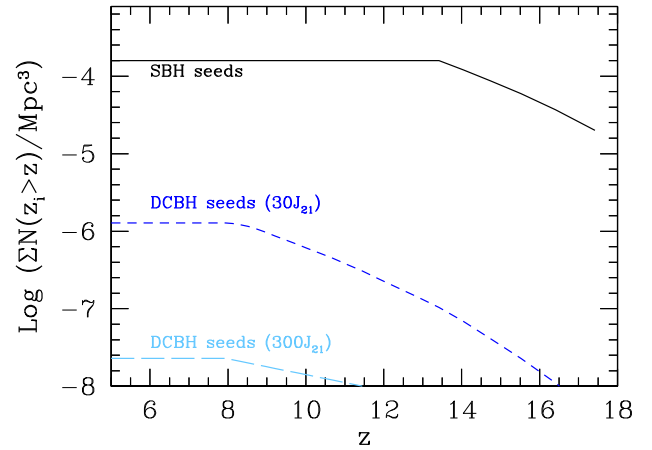


Figure 1. The cumulative redshift distribution of the number density of light (stellar BH; solid black line) and heavy (DCBH) seeds in our model. For the latter, we show the distribution of DCBH seeds using a value of $\alpha = 30$ (short dashed blue line) and $\alpha = 300$ (long-dashed light-blue line).

entirely in reproducing the observed active galactic nuclei (AGNs) luminosity function at $z = 5$ – 6 , given their extremely low number densities; consider the long-dashed cyan line in Fig. 1 that has a value of about $10^{-7.6} \text{ Mpc}^{-3}$ and compare this to an observed AGN number density of $\sim 10^{-5.5}$ – $10^{-4.5} \text{ Mpc}^{-3}$ (Kulkarni, Worseck & Hennawi 2018; Vito et al. 2018). Models with $\alpha = 30$ are nearly compatible with the faint end of the observed luminosity function at $z = 5$, but fail to produce enough AGNs at $z = 6$ unless the seed mass is above $10^4 M_\odot$. In summary, both ‘light’ and ‘heavy’ seed models have uncertainties and problems associated with their formation and growth. Given this state-of-the-art in the field of BH formation, our model comprehensively explores the allowed parameter space, includes a realistic approach to the dependence of BH growth on the host mass, and presents a thorough comparison with observations in Section 3 in order to explore the consequences of current uncertainties.

2.2 Star formation and supernova feedback

A newly formed stellar population of mass $M_*(z)$ at redshift z can impart the interstellar medium (ISM) with a total SNII energy E_{SN} given by

$$E_{\text{SN}} = f_*^w E_{51} \nu M_*(z) \equiv f_*^w v_s^2 M_*(z). \quad (1)$$

Here, each SNII is assumed to impart an (instantaneous) explosion energy of $E_{51} = 10^{51} \text{ erg}$ to the ISM and $\nu = [134 M_\odot]^{-1}$ is the number of SNII per stellar mass formed for a Salpeter IMF between 0.1 and $100 M_\odot$; we maintain this IMF throughout this work. The values of E_{51} and ν yield $v_s = 611 \text{ km s}^{-1}$. Finally, f_*^w is the fraction of the SN explosion energy that couples to gas.

For any given halo, the energy E_{ej} required to unbind and eject the ISM gas not converted into stars can be expressed as

$$E_{\text{ej}}(z) = \frac{1}{2} [M_{\text{gi}}(z) - M_*(z)] v_e^2, \quad (2)$$

where $M_{\text{gi}}(z)$ is the *initial* gas mass in the galaxy, prior to any star formation or BH accretion, at epoch z . Further, the escape velocity v_e can be expressed in terms of the halo rotational velocity, v_c , as $v_e = \sqrt{2} v_c$.

We then define the *ejection efficiency*, f_*^{ej} , as the fraction of gas that must be converted into stars to ‘blow away’ the remaining gas

from the galaxy (i.e. $E_{\text{ej}} \leq E_{\text{SN}}$). This can be calculated by imposing $E_{\text{ej}} = E_{\text{SN}}$, leading to

$$f_*^{\text{ej}}(z) = \frac{v_c^2(z)}{v_c^2(z) + f_*^w v_s^2}. \quad (3)$$

The *effective star formation efficiency* for any halo is then expressed as

$$f_*^{\text{eff}} = \min[f_*, f_*^{\text{ej}}], \quad (4)$$

where f_* is a free parameter representing the maximum instantaneous star formation efficiency – this parameter is fixed by matching to the bright end of the observed LBG UVLF as explained in Section 3.1.

In this formalism, the newly formed stellar mass formed at z can be expressed as

$$M_*(z) = M_{\text{gi}}(z) f_*^{\text{eff}}. \quad (5)$$

In the spirit of maintaining simplicity, we assume that every stellar population has a fixed metallicity of $0.05 Z_\odot$ and each newly formed stellar population has an age of 2 Myr. Using these parameters with the population synthesis code STARBURST99 (Leitherer et al. 1999), the rest-frame UV luminosity (between 1250 and 1500 Å) from a newly formed stellar mass can be expressed as

$$L_*^{\text{UV}} = 10^{33.077} \left(\frac{M_*}{M_\odot} \right) \text{erg s}^{-1} \text{Å}^{-1}. \quad (6)$$

This star formation episode must then result in a certain amount of gas, $M_*^{\text{ge}}(z)$, being ejected from the galaxy at the given z -step. The value of $M_*^{\text{ge}}(z)$ depends on whether $f_*^{\text{eff}} = f_*$ or $f_*^{\text{eff}} = f_*^{\text{ej}}$: while the galaxy is an ‘efficient star former’ in the former case, which can support new stellar mass being formed without losing much of its gas, the latter case is true for a ‘feedback-limited’ system that loses all of its ISM gas after star formation. Mathematically, M_*^{ge} can be calculated as

$$M_*^{\text{ge}}(z) = [M_{\text{gi}}(z) - M_*(z)] \frac{f_*^{\text{eff}}}{f_*^{\text{ej}}}. \quad (7)$$

The final gas mass, $M_*^{\text{gf}}(z)$, remaining in the galaxy at that redshift step, after star formation and SN feedback, can then be expressed as

$$M_*^{\text{gf}}(z) = [M_{\text{gi}}(z) - M_*(z)] \left[1 - \frac{f_*^{\text{eff}}}{f_*^{\text{ej}}} \right]. \quad (8)$$

2.3 BH growth and feedback

Once seeded, BHs can grow via accretion and mergers. We discuss BH growth via accretion in this section and the merger-driven growth is deferred to Section 2.5 that follows. At any given redshift, the Eddington mass accretion rate, \dot{M}_{ed} , for a BH of mass M_{bh} can be calculated as

$$\dot{M}_{\text{ed}}(z) = \frac{4\pi G M_{\text{bh}}(z) m_p}{\sigma_T \epsilon_r c}, \quad (9)$$

where G is the gravitational constant, m_p is the proton mass, σ_T is the Thomson scattering optical depth, ϵ_r is the BH radiative efficiency, and c is the speed of light. Given our merger tree time-step of $\Delta t = 20$ Myr, the total mass that can be accreted at the Eddington rate in one time-step is $M_{\text{ed}}(z) = (1 - \epsilon_r) \dot{M}_{\text{ed}}(z) \times \Delta t$.

Further, the gas mass accreted by the BH in a given time-step, $M_{\text{bh}}^{\text{ac}}(z)$, is calculated as

$$M_{\text{bh}}^{\text{ac}}(z) = \begin{cases} f_{\text{ed}} M_{\text{ed}}(z), & \text{if } x_r f_{\text{bh}}^{\text{ac}} M_*^{\text{gf}}(z) > f_{\text{ed}} M_{\text{ed}}(z) \\ x_r f_{\text{bh}}^{\text{ac}} M_*^{\text{gf}}(z), & \text{if } x_r f_{\text{bh}}^{\text{ac}} M_*^{\text{gf}}(z) < f_{\text{ed}} M_{\text{ed}}(z) \end{cases}, \quad (10)$$

where $x_r = (1 - \epsilon_r)$. Further, $f_{\text{bh}}^{\text{ac}}$ is the fraction of the available gas mass left, after star formation and supernova feedback, that can be accreted by the BH and f_{ed} , a free parameter, is the fractional Eddington rate of accretion. Using this formalism, the BH accretes either at a fraction of the Eddington rate or a fraction of the available gas mass, whichever is lower, i.e. $M_{\text{bh}}^{\text{ac}}(z) = \min[x_r f_{\text{bh}}^{\text{ac}} M_*^{\text{gf}}(z), f_{\text{ed}} M_{\text{ed}}(z)]$. Matching to the AGN UVLF (Section 3.1) requires $f_{\text{ed}} = 7.5 \times 10^{-5}$ below a critical halo mass of $M_{\text{bh}}^{\text{crit}} = 10^{11.25} [\Omega_m(1+z)^3 + \Omega_\Lambda]^{-0.125}$ (see also Bower et al. 2017) and $f_{\text{ed}} = 1$ above this mass at any redshift. This accretion will yield a BH feedback energy of

$$E_{\text{bh}} = f_{\text{bh}}^w \epsilon_r M_{\text{bh}}^{\text{ac}}(z) c^2, \quad (11)$$

where f_{bh}^w is the efficiency of BH feedback coupling to the gas.

The BH feedback required to eject the leftover gas (after accretion) can be expressed as

$$E_{\text{bh}}^{\text{ej}}(z) = \frac{1}{2} [M_*^{\text{gf}}(z) - M_{\text{bh}}^{\text{ac}}(z)] v_c^2. \quad (12)$$

The *effective BH feedback* is therefore taken to be the minimum between the energy required to eject all the gas up to the maximum value such that $E_{\text{bh}}^{\text{eff}} = \min[E_{\text{bh}}, E_{\text{bh}}^{\text{ej}}]$. The final gas mass left in the halo, which can be carried over for mergers, after BH accretion and feedback can then be calculated as

$$M_{\text{bh}}^{\text{gf}}(z) = [M_*^{\text{gf}}(z) - M_{\text{bh}}^{\text{ac}}(z)] \left[1 - \frac{E_{\text{bh}}^{\text{eff}}}{E_{\text{bh}}^{\text{ej}}} \right]. \quad (13)$$

Using the above formalism, the total luminosity produced by the BH can be expressed as

$$L_{\text{bh}} = \frac{\epsilon_r M_{\text{bh}}^{\text{ac}}(z) c^2}{\Delta t} [L_\odot]. \quad (14)$$

This is converted into the B -band luminosity using the results of Marconi et al. (2004) where $\log(L_{\text{bh}}/v_B L_{\nu B}) = 0.80 - 0.067(\log L_{\text{bh}} - 12) + 0.017(\log L_{\text{bh}} - 12)^2 - 0.0023(\log L_{\text{bh}} - 12)^3$. Finally, we use $L_\nu \propto \nu^{-0.44}$ to convert this B -band luminosity into the BH UV luminosity $L_{\text{bh}}^{\text{UV}}$. In order to build the AGN UVLF, we also account for AGN obscuration by multiplying the number density of BHs of a given luminosity by the correction factors proposed in Ueda et al. (2014).

2.4 Smooth accretion from the intergalactic medium (IGM)

The merger of haloes is accompanied by ‘smooth accretion’ of dark matter from the IGM. In the analytic merger tree, this smoothly accreted dark matter mass is calculated as

$$M_{\text{dm}}^{\text{sa}}(z) = M_{\text{h}}(z) - \sum_{i=1}^N M_{\text{h}}(z + \Delta z), \quad (15)$$

where $M_{\text{h}}(z)$ is the halo mass at z and the second term on the RHS denotes the sum of the halo masses of all its progenitors at the previous redshift step; $N = 2$ in the case of the binary merger tree used in our study.

We make the reasonable assumption that the smooth accretion of dark matter is accompanied by the accretion of a cosmological ratio

of gas from the IGM such that the smoothly accreted gas mass M_g^{sa} at z can be written as

$$M_g^{\text{sa}}(z) = \frac{\Omega_b}{\Omega_m} M_{\text{dm}}^{\text{sa}}(z). \quad (16)$$

2.5 Merging galaxies and BHs

As galaxies merge, in addition to dark matter, they bring in stellar and gas mass with the latter depending on the star formation and BH accretion efficiencies of the progenitor haloes. As detailed in Sections 2.2 and 2.3, galaxies forming stars and/or accreting at the limit of BH feedback will only bring stellar mass into their successors resulting in ‘dry mergers’. On the other hand, haloes bringing in both stellar and gas mass result in ‘wet mergers’. Mergers of galaxies result in a summation of their dark matter, baryonic, and BH masses and total UV luminosities such that

$$M_{\text{dm}}(z) = \sum_{i=1}^N M_{\text{h}}(z + \Delta z) + M_{\text{dm}}^{\text{sa}}(z) \quad (17)$$

$$M_g(z) = \sum_{i=1}^N M_{\text{bh}}^{\text{gf}}(z + \Delta z) + M_g^{\text{sa}}(z) \quad (18)$$

$$M_{*,\text{tot}}(z) = M_*(z) + \sum_{i=1}^N M_*(z + \Delta z) \quad (19)$$

$$M_{\text{bh,tot}}(z) = M_{\text{bh}}(z) + \sum_{i=1}^N M_{\text{bh}}(z + \Delta z) \quad (20)$$

$$L_{\text{tot}}^{\text{UV}}(z) = L_*^{\text{UV}}(z) + L_{\text{bh}}^{\text{UV}}(z) + \sum L_*^{\text{UV}}(z + \Delta z). \quad (21)$$

Using STARBURST99, we find that the UV luminosity for a burst of stars (normalized to a mass of $1 M_\odot$ and metallicity $0.05 Z_\odot$) decreases with time as

$$\log \left(\frac{L_*^{\text{UV}}(t)}{\text{erg s}^{-1} \text{Å}^{-1}} \right) = 33.0771 - 1.33 \log(t/t_0) + 0.462, \quad (22)$$

where t is the age of the stellar population (in yr) at z and $\log(t_0/\text{yr}) = 6.301$. Finally, we make the limiting assumption that BH luminosity decays away within the time-step of 20 Myr, i.e. BH luminosity is only relevant in the redshift step in which the BH accretes.

We also explore two scenarios for the *time-scales* of galaxy mergers: *in the first*, halo mergers are accompanied by the mergers of galaxies and their constituent BHs. This *instantaneous merging* scenario sets the *upper limit* for the BH merger rate. *In the second scenario*, we include the fact that galaxies (and their BHs) merge after a ‘merging’ time-scale that can be calculated as (Lacey & Cole 1993)

$$\tau = f_{\text{df}} \Theta_{\text{orbit}} \tau_{\text{dyn}} \frac{M_{\text{host}}}{M_{\text{sat}}} \frac{0.3722}{\ln(M_{\text{host}}/M_{\text{sat}})}, \quad (23)$$

where M_{host} is the mass of the host including all the satellites, M_{sat} is the mass of the merging satellite, τ_{dyn} represents the dynamical time-scale, and f_{df} represents the efficiency of tidal stripping; $f_{\text{df}} > 1$ if tidal stripping is very efficient. For this work, we use $f_{\text{df}} = 1$. Further,

$$\Theta_{\text{orbit}} = \left(\frac{J}{J_c} \right)^{0.78} \left(\frac{r_c}{R_{\text{vir}}} \right)^2, \quad (24)$$

where J is the satellite’s specific angular momentum and J_c that of a satellite carrying the same energy and orbiting on a circular orbit. The last term represents the ratio between the circular radius (the radius of a circular orbit with the same energy) and the virial radius of the host. Θ_{orbit} is well modelled by a lognormal distribution such that $\log(\Theta_{\text{orbit}}) = -0.14 \pm 0.26$ (Cole et al. 2000); we randomly sample values from this distribution for each merger. Finally, the dynamical time-scale can be calculated as $\tau_{\text{dyn}} = \pi R_{\text{vir}}(z) V_{\text{vir}}(z)^{-1} = 0.1 \pi t_{\text{H}}(z)$, where R_{vir} and V_{vir} are the virial radius and velocity at z , respectively. We also make the limiting assumption that satellite galaxies, waiting to merge, neither form stars nor have any BH accretion of gas.

We show results from both scenarios in this work in order to bracket the uncertainty on our understanding of galaxy merger time-scales. In reality, however, the second scenario still gives a lower limit to the merger time of BH binaries: additional delays are to be expected once the binary forms after the dynamical friction time-scale. The binary has to harden further to reach the separation where GW emission becomes the dominant source of energy and momentum losses to bring the binary to coalescence and merge within the Hubble time. This additional delay is largely unconstrained and can range from tens to billions of years (for a review, see e.g. Barack et al. 2018). Additionally, for low-mass BHs dynamical friction could be ineffective during the galaxy merger phase since the deceleration of dynamical friction is proportional to the infalling BH mass resulting in the time-scale for orbital decay being inversely proportional to mass.

2.6 The impact of reionization

In this work, we also include the effects of the ultraviolet background (UVB) created during reionization that, by heating the ionized IGM to $T \sim 10^4$ K, can have an impact on the baryonic content of low-mass haloes (for a recent review, see e.g. Dayal & Ferrara 2018). Given that self-consistently calculating the impact of the UVB on galaxy formation remains an unsolved problem, we consider two scenarios: *the first* is in which UVB has no impact on the baryonic content of any halo. *The second* scenario is the one considering maximal UVB feedback in which the gas mass is completely photoevaporated for all haloes below a characteristic virial velocity of $V_{\text{vir}} = 40 \text{ km s}^{-1}$. Critically, while leaving the *number* of mergers unchanged, limiting the gas mass available for accretion on to BHs, the latter scenario results in a *lower limit* on the mass of the merging BH.

We run the above model for eight different scenarios (detailed in Table 1) that explore three key physical effects: (i) the impact of the LW background amplitude in calculating DCBH hosts (Section 2.1); (ii) the impact of instantaneous versus delayed mergers of galaxies and BHs (Section 2.5); and (iii) the impact of UV feedback on early galaxies and BHs (Section 2.6). In what follows, the model *insl*, which assumes the lowest LW amplitude for DCBHs, instantaneous galaxy mergers, and no UV feedback, is denoted as the *fiducial* model and provides the upper limit on our results. On the other hand, with the highest LW amplitude for DCBHs, delayed galaxy (and BH) mergers, and maximal UV feedback, the model *tdf4* provides the lower limit to our results.

3 COMPARING THEORETICAL GALAXY AND BH PROPERTIES TO OBSERVATIONS

Now that the theoretical model has been established, we compare model results to a number of observational data sets, for both

Table 1. Values of the free parameters used for the models are indicated in column 1. See the text in Section 2 for details.

Model	α	UVB	f_*	f_*^w	f_{bh}^w	$f_{\text{bh}}^{\text{ac}}$	$f_{\text{ed}}(M_{\text{h}} < M_{\text{bh}}^{\text{crit}})$	$f_{\text{ed}}(M_{\text{h}} > M_{\text{bh}}^{\text{crit}})$	ϵ_{r}
ins1	30	No	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
ins2	300	No	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
ins3	30	Yes	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
ins4	300	Yes	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
tdf1	30	No	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
tdf2	300	No	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
tdf3	30	Yes	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1
tdf4	300	Yes	0.02	0.1	0.003	5.5×10^{-4}	7.5×10^{-5}	1	0.1

galaxies and AGNs, as detailed in what follows. In terms of galaxies, we use the data sets accumulated for high- z LBGs, detected through a drop in luminosity blue-ward of the Lyman limit at 912 Å. The past few years have seen an enormous increase in LBG data due to a combination of state-of-the-art instruments such as (the Wide Field Came 3 onboard) the *Hubble Space Telescope* as well as refined selection techniques (e.g. Steidel et al. 1999). As for AGNs, a number of surveys, including the Sloan Digital Sky Survey and the Canadian–French high- z quasar surveys, and observations with the Subaru telescope, have yielded a statistical sample of AGN/QSO candidates at redshifts as high as $z \simeq 6$. In what follows, we compare four models that bracket the physically plausible range explored in this work (*ins1*, *ins4*, *tdf1*, and *tdf4*), detailed in Table 1, with a number of data sets, including the UVLFs, the SMD, the BH mass function, and the BH mass–stellar mass relation.

We note that, given their low number densities, both the ‘light’ and ‘heavy’ DCBH seeding cases yield very similar results for all the observational data sets discussed. For this reason, we limit our results to the ‘light DCBH seed’ case in this section.

3.1 The observed UVLF for star formation and BHs

The observed UVLF (number density of galaxies as a function of the absolute magnitude) and its redshift evolution offer one of the most robust tests of theoretical models of galaxy formation. We start by calculating the UV magnitudes, separately for star formation and AGN activity, for each theoretical galaxy and computing the associated UVLFs, as shown in Fig. 2. We start by discussing the LBG UVLF: first, matching to the bright end of the evolving UVLF requires a maximum star formation efficiency value of $f_* \simeq 2$ per cent. Secondly, we find that the fiducial model (*ins1*) is in excellent agreement with available LBG observations, ranging between $-22 \lesssim M_{\text{UV}} \lesssim -13$, at all $z \sim 5$ –10 as already shown in our previous works (e.g. Dayal et al. 2014). The inclusion of a delay in galaxy mergers (*tdf1*) has no sensible impact on the faint end of the UVLF – this is due to the fact that the progenitors of these low-mass haloes are SN feedback limited and hence do not bring in any gas while merging (dry mergers) as already pointed out previously (Dayal et al. 2014). On the other hand, the delay in galaxy mergers leads to an increasing reduction in the gas masses of higher mass haloes whose progenitors are not SN feedback limited and bring in gas in mergers (wet mergers), leading to a slightly steeper bright end. Finally, including the (maximal) impact of reionization feedback (*ins4* and *tdf4*), which photoevaporates the baryonic content of all galaxies with $V_{\text{vir}} \lesssim 40 \text{ km s}^{-1}$, only affects the faint end of the UVLF and leads to a cut-off at brighter magnitudes (from $M_{\text{UV}} \sim -14$ to -15) as compared to the continued rise excluding this effect (e.g. in models *ins1* and *tdf1*). In this work, models *ins1* and *tdf4*, therefore, bracket the plausible UVLF range. However,

it must be cautioned that the theoretical LBG UVLF has, so far, ignored the impact of dust enrichment, which is expected to have a relevant effect in decreasing the luminosities at the bright end.

Focusing on the AGN UVLF, the BH-powered UVLFs for all the four models discussed above are found to be in excellent agreement with all available AGN data at $z \sim 5$ and 6 as shown in Fig. 2. We start by noting that given the large masses ($M_{\text{h}} \gtrsim 10^{11.5} M_{\odot}$) associated with AGN/QSO host haloes, the BH UVLF is only relevant at $M_{\text{UV}} \lesssim -21$, corresponding to number densities $\lesssim 10^{-5} [\text{dex}^{-1} \text{Mpc}^{-3}]$ at $z \sim 5$ and 6. These results are in qualitative agreement with those of Ono et al. (2018), who find 100 per cent of the UV luminosity to come solely from stars for galaxies with $M_{\text{UV}} \gtrsim (-23)$ – (-24) . However, given that the AGN number densities are suppressed due to obscuration (see Section 2.3), calculating the fraction of galaxies dominated by AGNs requires a more thorough examination, which we defer to a future work. Finally, we note that the contribution of BH-powered luminosity could be one explanation for observed UVLFs that are shallower than the exponentially declining Schechter function at these high- z (e.g. Ono et al. 2018).

We find that the AGN UVLF is extremely similar for heavy BH seeds with α varying over an order of magnitude (from 30 to 300) for the four models discussed above. This is probably to be expected given the extremely low number of heavy BH seeds as compared to the number of light BH seeds as pointed out in Section 2.1; the latter therefore clearly dominate the UVLF. As for the merger time-scales, including a delay in the mergers of galaxies (and BHs) results in a smaller BH growth. This is reflected in a lower final BH mass in a given halo (also see Section 3.3 that follows). However, this only leads to minor changes in the UVLF that are indistinguishable within the scatter shown by the four models considered here. Further, given the large masses of AGN hosts, reionization feedback has no relevant effect on the AGN UVLF. Finally, looking at the redshift evolution of the AGN UVLF, we find that it shows a sharper redshift evolution compared to the star formation-powered UVLF, given the increasing paucity of their high-mass hosts. To quantify this effect, let us focus on a magnitude of $M_{\text{UV}} = -20$: while the star formation-driven UVLF only evolves by a factor of 3 between $z \sim 5$ and 7, the AGN UVLF (negatively) evolves by roughly three orders of magnitude over the same redshift range.

3.2 The LBG SMD

We now compare the theoretical SMD to that observationally inferred for LBGs. We start by comparing to observed LBGs with $M_{\text{UV}} \lesssim -17.7$ as shown in Fig. 3. As seen, while all the four models (*ins1*, *ins4*, *tdf1*, and *tdf4*) are in excellent agreement with the data they are offset in normalization from each other while following

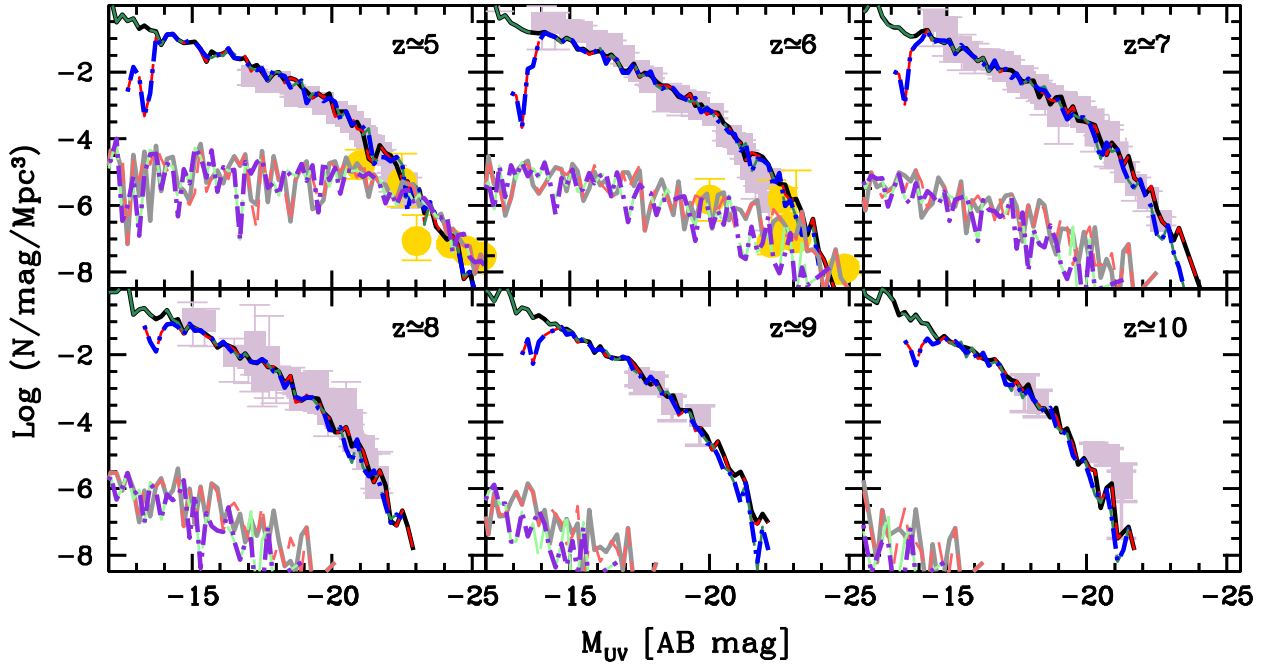


Figure 2. The UVLF from $z \simeq 5$ to 10 as marked in the panels. In each panel, the violet points show the available LBG data collected both using space- and ground-based observatories at: (a) $z \simeq 5$ (Bouwens et al. 2007; McLure et al. 2009); (b) $z \simeq 6$ (McLure et al. 2009; Bouwens et al. 2015; Livermore, Finkelstein & Lotz 2017); (c) $z \simeq 7$ (Bouwens et al. 2010; Castellano et al. 2010; McLure et al. 2010, 2013; Bowler et al. 2015; Atek et al. 2015; Livermore et al. 2017); (d) $z \simeq 8$ (Bouwens et al. 2010; McLure et al. 2010; Bradley et al. 2012; McLure et al. 2013; Atek et al. 2015; Livermore et al. 2017); (e) $z \simeq 9$ (McLure et al. 2013; Oesch et al. 2013); and (f) $z \simeq 10$ (Oesch et al. 2014; Bouwens et al. 2015). In each panel, the yellow points show the AGN data collected at $z \sim 5$ (McGreer et al. 2013; Parsa, Dunlop & McLure 2018) and $z \sim 6$ (Willott et al. 2010; Kashikawa et al. 2015; Jiang et al. 2016; Parsa et al. 2018). In each panel, lines show model UVLFs for galaxies and BHs for the following models summarized in Table 1, which bracket the range of UVLFs allowed in the presence/absence of a UVB and for both instantaneous and delayed (by a merging time-scale) mergers: *ins1* (galaxies: solid black line; BH: solid grey line), *ins4* (galaxies: short dashed red line; BH: short dashed light-red line), *tdf1* (galaxies: long dashed green line; BH: long-dashed light green line), and *tdf4* (galaxies: dot-dashed blue line; BH: dot-dashed purple line).

very similar slopes such that $\text{SMD} \propto (1+z)^{0.42}$. As might be expected, model *ins1* provides the upper limit to the SMD results for observed galaxies. Including the effects of delayed galaxy mergers (*tdf1*) results in a small decrease in the SMD values by about 0.1 dex. However, assuming instantaneous mergers while including maximal UVB suppression (*ins4*) only results in an SMD that is different from the fiducial case by a negligible 0.03 dex. These results clearly imply that a delay in the merger time-scales is more important than the effect of a UVB for these high-mass systems. Finally, the lower limit to the SMD results is provided by model *tdf4* that is about 0.13 dex lower than the fiducial results. These slight changes in the SMD normalization show that most of the stellar mass is assembled in massive progenitors (see also Dayal et al. 2013) with low-mass progenitors – that merge after a dynamical time-scale (*tdf1*), are reionization suppressed (*ins4*), or include both these effects (*tdf4*) – contributing only a few per cent to the stellar mass for observed galaxies.

On the other hand, the impact of reionization feedback and a delay in the merging time-scale are much more dramatic when considering the entire galaxy population (without any limiting magnitudes used). In this case, the fiducial model, *ins1*, shows a slope that evolves with redshift as $\text{SMD} \propto (1+z)^{0.24}$. Given that in this case the SMD is dominated by the contribution from low-mass haloes, the situation flips as compared to that discussed above: the merger time-scale has a negligible effect on the SMD of all galaxies and shows essentially the same amplitude and slope as the fiducial case. However, the UVB suppression of the gas mass of low-mass haloes results in

both a decrease in the amplitude (by about 0.23 dex) and a, more dramatic, steepening of the SMD slope such that $\text{SMD} \propto (1+z)^{0.31}$ for models *ins4* and *tdf4*.

3.3 The BHMF and occupation fraction

We now discuss the BHMF that expresses the number density of BHs as a function of their mass, the results of which at $z \simeq 6$ are shown in Fig. 4. As expected, the number density of BHs increases with decreasing BH mass as shown in the figure. The observed BHMF at $z \sim 6$ extends from $M_{\text{bh}} \sim 10^{7-10} M_{\odot}$. Our theoretical results for all the four models discussed above are in good agreement with the data within error bars as seen in the same figure. Naturally, the fiducial model (*ins1*), extending from $M_{\text{bh}} \sim 10^{4.8-8.8} M_{\odot}$, yields the upper limit to the BHMF. Including a delay in the merger times for BHs (*tdf1*) leads to a decrease in the maximum mass attained by the BHs ($M_{\text{max}} \sim 10^8 M_{\odot}$), showing that gas brought in by merging progenitors haloes has a significant contribution to the growth of these high-mass systems. On the other hand, reionization feedback alone (*ins4*) has a negligible effect on the growth of high-mass haloes (as discussed in Section 3.2 above), yielding a BHMF in close agreement with the fiducial one. Finally, the model *tdf4*, including both the impact of delayed mergers and the UVB, yields results quite similar to *tdf1* and provides the lower limit to the BHMF. We recall that our model is not aimed at (re)producing rare luminous quasars powered by very massive BHs (see Valiante et al. 2016; Pezzulli, Valiante & Schneider 2016, and references therein for

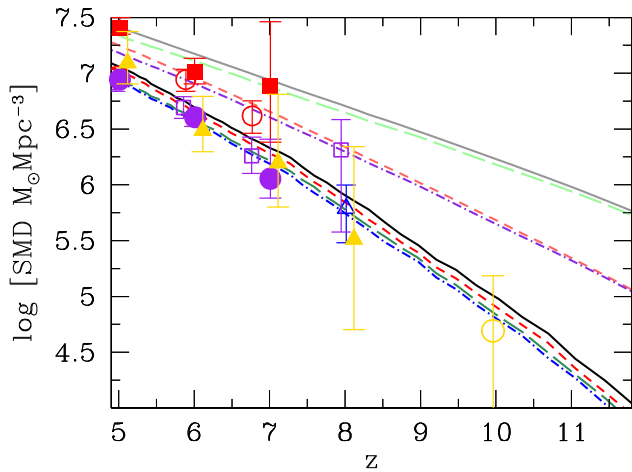


Figure 3. The LBG SMD as a function of redshift. The points show the observational data collected by: González et al. (2011, red open circles), Labbé et al. (2013, blue open triangles), Stark et al. (2013, purple open squares), Oesch et al. (2014, yellow open circles), Duncan et al. (2014, red filled squares), Grazian et al. (2015, purple filled circles), and Song et al. (2016, yellow filled triangles). We show results for galaxies with $M_{UV} < -17.7$, which can be directly compared to observational data points for the following models shown in Table 1: *ins1* (solid black line), *ins4* (dot-dashed red line), *tdf1* (long dashed green line), and *tdf4* (dot-dashed blue line). We also show results for the total SMD obtained by summing over all galaxies at a specific z for the same models noted above: *ins1* (solid grey line), *ins4* (dot-dashed light red line), *tdf1* (long dashed light green line), and *tdf4* (dot-dashed purple line).

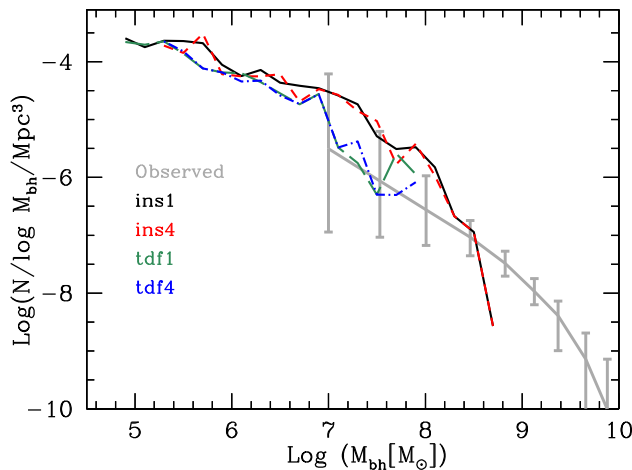


Figure 4. The BHMf at $z \sim 6$. We compare observational results (grey line with error bars) from Willott et al. (2010) to those from our models bracketing the plausible physical range: *ins1* (solid black line), *ins4* (short dashed red line), *tdf1* (long-dashed green line) and *tdf4* (dot-dashed blue line). As shown, the shape of the BHMf is independent of the inclusion of UV feedback and the merger time-scales used. However, the final BH masses are naturally lower when including a delay in the merger time-scales as opposed to instantaneous mergers.

models focused on the most massive haloes and BHs) but at the bulk of the population of massive BHs. It should therefore not be surprising that the BH mass function does not extend to the highest BH masses observed.

We also show the BH occupation fraction in Fig. 5. As shown, galaxies with a halo mass $M_h \gtrsim 10^{10.2}$ have an occupation fraction of

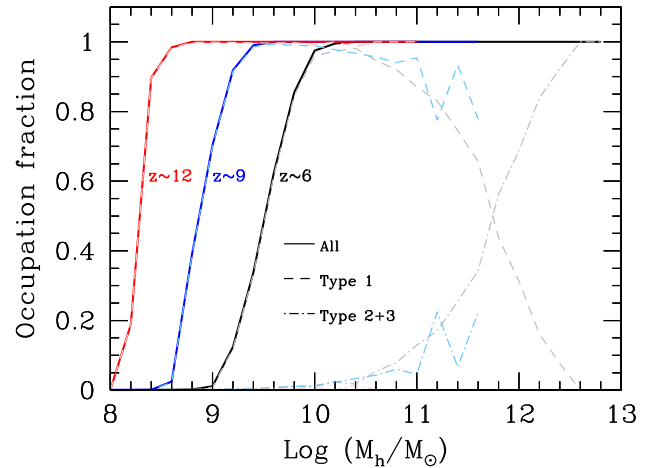


Figure 5. The BH occupation fraction as a function of halo mass for $z \sim 6$ (black lines), $z \sim 9$ (blue lines), and $z \sim 12$ (red lines) as marked. The solid lines show the occupation fraction for all BHs; the short dashed and dot-dashed lines show results for Type 1 (stellar BHs) and Type 2 + 3 (DCBHs), respectively.

1 by $z \sim 6$. As expected, most of these are stellar BHs except DCBHs that dominate for the most massive haloes. The BH occupation fraction also shifts to progressively lower masses with increasing redshift. This is because of two reasons: first, in our model, only starting leaves above $z = 13$ are seeded with BHs; the increasing number of starting leaves forming at lower redshifts are devoid of any BHs. Secondly, low-mass haloes continually increase in mass with decreasing redshift. We note that our results are qualitatively in good agreement with those obtained from previous works (e.g. Tanaka & Haiman 2009). Finally, we stress that the enhancement of the LW seen by any haloes only depends on its bias at that redshift and we have ignored the impact of clustered sources that could enhance the LW intensity seen by haloes in overdense environments. Our results must therefore be treated as a lower limit on the DCBH number density and, hence, the Type 2 + 3 occupation fraction.

3.4 The BH mass–stellar mass relation

Constraints on the relation between BHs and galaxies at high redshift are scant. In general, since the only confirmed BHs at these redshifts are those powering powerful quasars, the stellar mass of the host cannot be measured (not to mention the stellar velocity dispersion or bulge mass) because the light from the quasar overshines the host galaxy. The best estimates of the host properties for these powerful quasars are obtained through measures of the cold (molecular) gas properties in sub-mm observations, where a dynamical mass, based on the velocity dispersion of the gas and the radius of the emitting region, can be measured (e.g. Venemans et al. 2016; Shao et al. 2017; Decarli et al. 2018, and references therein). For these quasars, the BH to dynamical mass is skewed to values much larger than the ratio of BH to stellar or bulge mass in the local Universe. As discussed in Volonteri & Stark (2011), there are reasons to believe that such high mass ratios should not characterize the whole BH population. Beyond the Malmquist bias causing a more frequent selection of overmassive BHs in low-mass hosts (Lauer et al. 2007; Salvander et al. 2007), only undermassive and low-accretion BHs can explain the lack of widespread AGN detections in LBGs. That BHs in low-mass galaxies are indeed expected to grow slowly and lag behind the

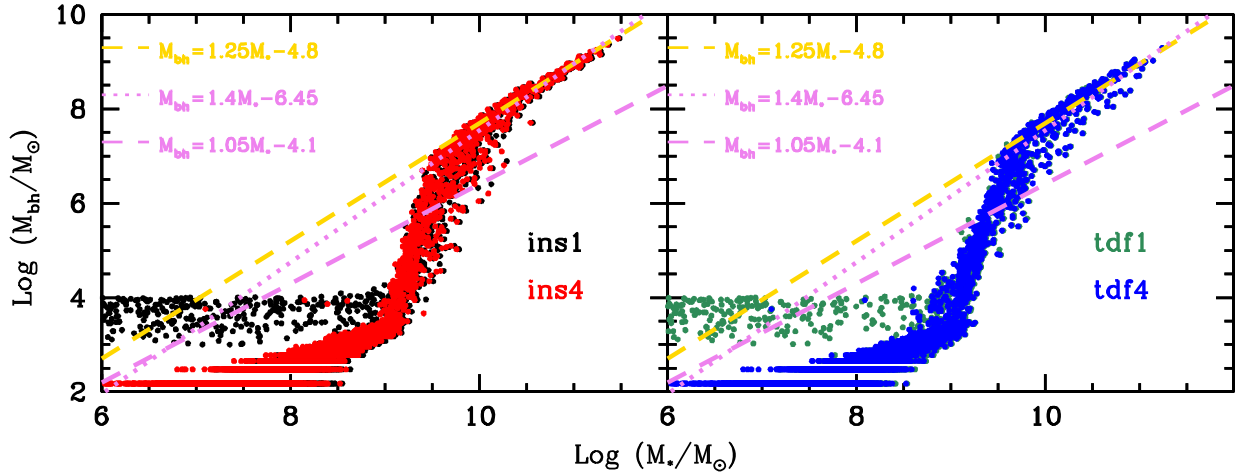


Figure 6. The BH mass–stellar mass relation for $z \simeq 5$ for two models that bracket the expected range: instantaneous mergers with/without UV feedback (*ins1* using black points and *ins4* using red points; left-hand panel) and mergers after a merging time-scale with/without UV feedback (*tdf1* using green points and *tdf4* using blue points; right-hand panel). In both the panels, we show two relations derived using galaxies in the nearby Universe: $M_{\text{bh}} = 1.4M_* - 6.45$ derived for high-stellar mass ellipticals and bulges and $M_{\text{bh}} = 1.05M_* - 4.1$ for moderate-luminosity AGNs in low-mass haloes (Volonteri & Reines 2016), as marked. In each panel, we also show the best-fitting relation from our model for high-stellar mass galaxies: $\text{Log} M_{\text{bh}} = 1.25M_* - 4.8$. As seen, our theoretical model yields a non-linear scaling such that BHs in low-mass galaxies are ‘stuck’ at their initial mass; the BH masses of high-mass hosts, on the other hand, are strongly correlated with the stellar mass and are in excellent agreement with the results derived for low- z high-stellar mass galaxies.

host has now been confirmed in many numerical investigations (Dubois et al. 2015; Anglés-Alcázar et al. 2017; Bower et al. 2017; Habouzit, Volonteri & Dubois 2017). Our implementation of BH growth includes a stunted growth in low-mass galaxies and we obtain a BH mass–stellar mass relation in agreement with numerical investigations, a non-linear scaling where BHs in low-mass galaxies are ‘stuck’ at their initial mass (Bower et al. 2017; Habouzit et al. 2017). BHs in high-mass hosts, on the other hand, can be above the $z = 0$ scaling, as shown in Fig. 6.

Quantitatively, we find that the BH mass–stellar mass relation is strongly correlated for high-stellar mass ($M_* \gtrsim 19^{9.5} M_\odot$) galaxies and is best expressed by the relation $M_{\text{bh}} = 1.25M_* - 4.8$ at $z \simeq 5$; the relation flattens below such masses. Including the impact of the UVB (*ins4*) has no impact on this relation at the bright end. However, the suppression of gas mass in low-mass haloes naturally results in lower BH masses by as much as two orders of magnitude for a given stellar mass. As noted above in Section 3.3, the inclusion of a delay in galaxy merging time-scales results in a decrease in the mass of the most massive BHs (by about 0.8 dex) as seen from the right-hand panel of the same figure although it has no impact on the high-mass slope. Further, the results from *ins4* and *tdf4* are quite similar as also expected from the discussion in Section 3.3 above, yielding the lower limit to the $M_{\text{bh}}-M_*$ relation. Finally, the best-fitting relation derived for high-stellar mass galaxies from our model is in excellent agreement with the relation $M_{\text{bh}} = 1.4M_* - 6.45$ derived for high-stellar mass ellipticals and bulges in the nearby Universe (Volonteri & Reines 2016).

4 LISA AND GWS FROM THE HIGH- z UNIVERSE

Now that we have shown the theoretical galaxy and BH properties to be in excellent agreement with observations, we can extend our calculations to the GWs expected from the mergers of such high- z BHs. In this work, any merger falls into one of the following three categories: (i) *type1 – stellar BH mergers*: mergers of two stellar

BH seeds; (ii) *type 2 – mixed mergers*: mergers of a stellar BH seed with a DCBH, and (iii) *type 3 – DCBH–DCBH mergers*: extremely rare, these are mergers of two DCBH seeds. In the last category, we also include mergers of a DCBH with a mixed merger in the past.

A system with two BHs revolving around each other forms an accelerated mass quadrupole that causes emission of GWs at the expenses of orbital energy with a catastrophic outcome: as the binary emits GWs, its semimajor axis shrinks (‘inspiral’ phase) until the two BHs merge and, after shedding any extra residual energy (‘ringdown’ phase), a newly born stationary BH forms. The GW signal increases in amplitude and frequency at an accelerated pace with the emission peaking at merger, i.e. roughly at the innermost stable circular orbit (ISCO). The peak frequency at the ISCO for a non-spinning BH can be expressed as twice

$$f_{\text{ISCO}} = \frac{1}{6\sqrt{6}(2\pi)} \frac{c^3}{GM(1+z)}, \quad (25)$$

where M is the total binary mass. Beyond the peak, the signal is exponentially damped. Massive BHs ($M_{\text{bh}} > 10^3 M_\odot$) at high redshifts emit at frequencies ($\ll 1$ Hz) much lower than the range of ground-based GW detectors. To detect massive BHs through GWs, much longer interferometric arms, of a million kilometres, are needed, which can be only realized in space. In this section, we forecast the detection performance of the space-based European Space Agency mission *LISA* for BH binaries in the early Universe ($z > 4$), in the evolutionary framework presented above. *LISA* is a space-based GW laser interferometer, proposed to be launched in 2034, that consists of three spacecrafts in an equilateral triangle constellation. The interferometer’s arms are proposed to be 2.5×10^6 km in length, resulting in an optimal frequency range between ≈ 1 mHz and 0.1 Hz (Fig. 7).

For each BH merger, the optimized value of the signal-to-noise ratio (SNR) associated with the wave model is calculated based on the matched-filtering technique. By assuming the noise to be

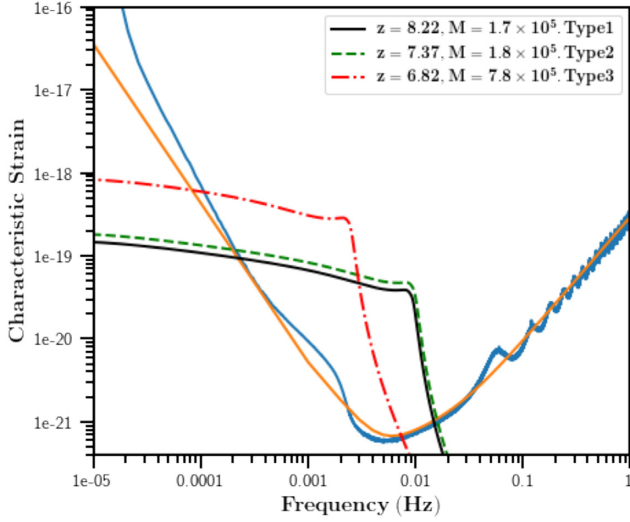


Figure 7. *LISA* sky-averaged dimensionless sensitivity curve ($\sqrt{f \times S_n}$) as a function of frequency for 4 yr of observation time: numerical calculation (Amaro-Seoane et al. 2017, blue line) and analytical approximation (yellow line; model A2N2 from Klein et al. 2016). The numerical curve accounts for the Galactic binary stochastic foreground noise that causes the ‘bump’ around 10^{-3} Hz. We also overplot the dimensionless characteristic strain $h_c = 2f\tilde{h}(f)$ for the averaged (in total mass and redshift) detected BH binaries in the three different types of mergers (using the ‘light’ DCBH seeds model) considered in this work. Note that the average detected binary is at an increasing redshift and has a decreasing total mass going from type 3 to type 1 mergers.

stationary and Gaussian with zero mean, the SNR is given by

$$\left(\frac{S}{N}\right)^2 = 4 \int_{f_{\min}}^{f_{\max}} \frac{|\tilde{h}(f)|^2}{S_n(f)} df, \quad (26)$$

where $|\tilde{h}(f)|$ is the amplitude of the GW signal in frequency domain and $S_n(f)$ is the noise power spectra density (PSD) function. Here, f_{\min} is the binary frequency when it is first observed (i.e. at $t = 0$) and f_{\max} is either its frequency at the end of the mission’s lifetime or at merger, whichever happens first. In the following sections, we detail the calculation of the signal [$h(f)$; Section 4.1] and the noise (S_n) and integration limits (Section 4.2). We end by presenting our results in Section 4.3.

4.1 The GW signal

GW detectors generally work with time-dependent scalars, $h(t)$, as their output. The scalar describes the changes in the detector after the passage of waves. In case of laser interferometers, $h(t)$ represents the phase shift of the laser beam (or equivalently, the change in the detector’s arm length) and can be expressed as

$$h(t) = F_{\times} h_{\times}(t) + F_{+} h_{+}(t), \quad (27)$$

where $h_{\times}(t)$ and $h_{+}(t)$ are the GW polarizations. Further, F_{\times} and F_{+} are the detector’s pattern functions that depend both on the properties of the detector and the position of the source in the sky. After averaging the signal over the sky position and transforming it into the frequency domain, we obtain

$$|\tilde{h}(f)|^2 = |\tilde{A}(f)|^2 \times |Q|^2, \quad (28)$$

where $\tilde{A}(f)$ is the wave amplitude in the frequency domain and Q is a geometrical factor containing information about the pattern

functions. For our choice of the detector’s configuration, we use $Q = \frac{2}{10}\sqrt{3}$, which accounts for averaging over sky position and binary inclination (see Amaro-Seoane et al. 2017).

The calculation of $\tilde{A}(f)$ should in principle be performed with a fully relativistic (NR) numerical code. However, such calculations are numerically expensive and, in fact, only necessary for modelling the highly relativistic end of the inspiral phase and merger. The inspiral phase, where orbital velocities are much lower than the speed of light, can instead be satisfactorily reproduced with an analytical post-Newtonian (PN) formalism. These considerations inspired the so-called ‘phenomenological models’ that give a complete analytical wave model by matching the PN and NR waveforms in the region where the PN approximation breaks down. To calculate $\tilde{A}(f)$, we model the waveform with the phenomenological model ‘PhenomC’, which has the advantage of producing the waveform directly in frequency domain, convenient for data-analysis applications (for a detailed description of the code, see Santamaría et al. 2010).

4.2 The instrumental and source noise

We numerically calculate the sky-averaged noise PSD, S_n , for *LISA* using the *LISA*-consortium simulator, which takes into account different instrumental noises as well as the stochastic background from unresolved Galactic binaries. The most notable contribution to the latter comes from Galactic white dwarf binaries that *LISA* is unable to resolve individually (e.g. Amaro-Seoane et al. 2012). The number of these sources is expected to decrease as the mission progresses and a larger number of foreground sources are detected and removed. *LISA*’s sky-averaged sensitivity curve ($\sqrt{f \times S_n}$) adopted in this paper for the SNR calculation corresponds to a 4-yr observing time and is presented (using the blue line) in Fig. 7. For convenience, the frequency limits of the integral (equation 26) are instead calculated adopting an analytical fit to *LISA*’s PSD of the form

$$S_n(f) = \frac{20}{3} \frac{4S_{n,\text{acc}}(f) + S_{n,\text{sn}}(f) + S_{n,\text{omn}}(f)}{L^2} \times \left[1 + \left(\frac{f}{\frac{0.41c}{2L}} \right)^2 \right], \quad (29)$$

(yellow solid line in Fig. 7) from Klein et al. (2016). In the above equation, L is the detector’s arm length. Further, $S_{n,\text{acc}}$, $S_{n,\text{sn}}$, and $S_{n,\text{omn}}$ are the noise components due to low-frequency acceleration, shot noise, and other measurement noise, respectively. Instead of performing a formal fit to the numerical curve in order to estimate the noise parameters, we adopt the following values from those reported in Klein et al. (2016):

$$\begin{aligned} S_{n,\text{acc}} &= \frac{9 \times 10^{-30}}{(2\pi f)^4} \left(1 + \frac{10^{-4}}{f} \right) [\text{m}^2 \text{Hz}^{-1}], \\ S_{n,\text{sn}} &= 2.22 \times 10^{-23} [\text{m}^2 \text{Hz}^{-1}], \\ S_{n,\text{omn}} &= 2.65 \times 10^{-23} [\text{m}^2 \text{Hz}^{-1}], \end{aligned} \quad (30)$$

corresponding to an $L = 2$ Mkm arm length (model A2N2L6; Klein et al. 2016). Both the numerical and analytical curves in Fig. 7 are calculated for the current *LISA* design, which presents three spacecrafts connected by six links.

A visual comparison between our analytical (equation 29) and numerical curves shows a sufficiently close match for our purposes for frequencies $> 10^{-4}$ Hz. We note that the analytical curve does not account for a stochastic background (the bump around 10^{-3} Hz). This omission allows low-mass BH binaries ($\sim 10^3 M_{\odot}$), which

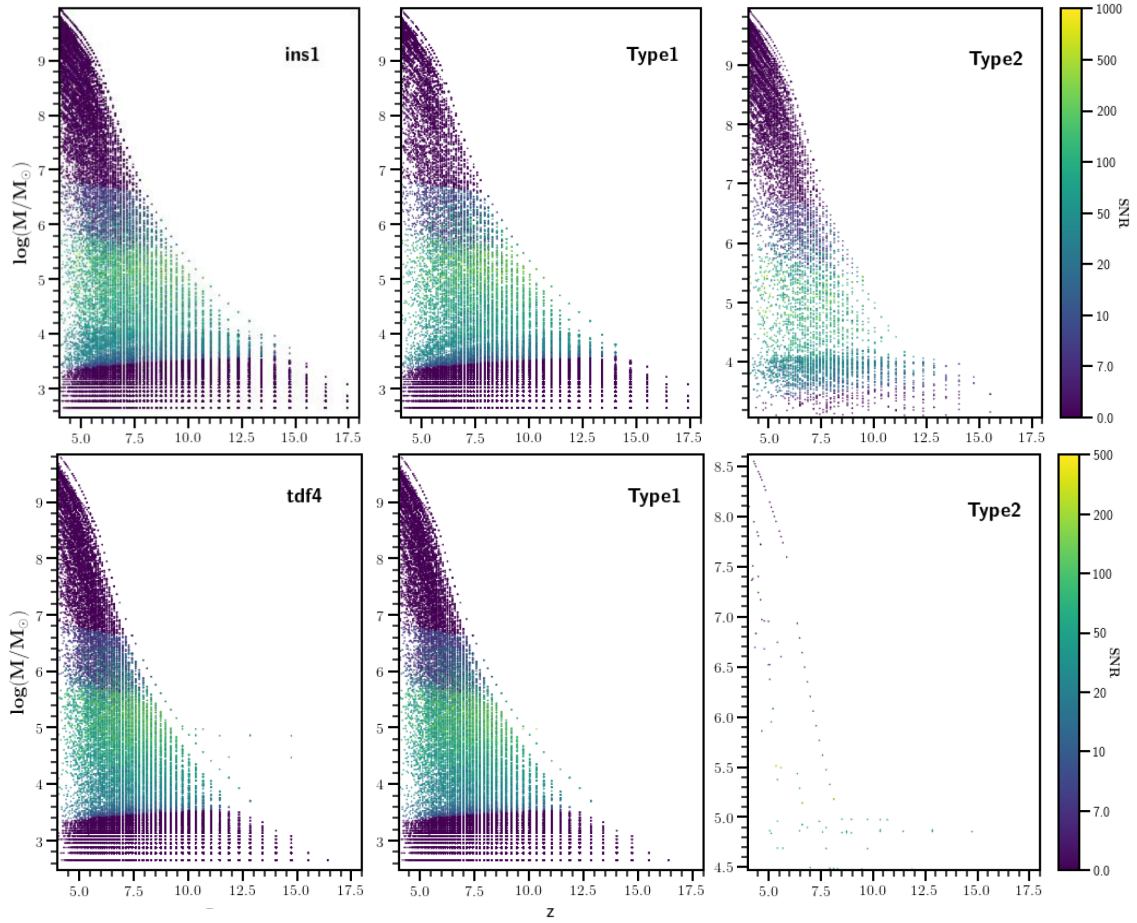


Figure 8. The SNR as a function of intrinsic total binary mass and z for a 4-yr mission duration. The columns from left to right show results for all mergers, type 1 mergers (mergers of two stellar BHs) and type 2 mergers (mergers of a stellar BH and a DCBH); there are no detections of type 3 BHs (mergers of two DCBHs). As marked, the upper and lower rows correspond to results for models *ins1* and *tdf4* with ‘light’ DCBH seeds, respectively. The *LISA* detectability window is such that binaries with $\text{SNR} > 7$ have redshifts between $z = 5$ and 13 and a total mass between $M \simeq 10^{3.5-5.6} M_\odot$ with the exact value depending on the model and merger type. The characteristic strain for the average binary is traced in Fig. 7.

cross the sensitivity curve near its bottom, to stay longer in the observed band. However, we will show that, even with this extra integration time, their SNR is never above the detection threshold. Finally, we overplot the signals from representative detected merger events. The examples considered in this figure are the average (in total mass and redshift) detected binary for each of the three types of mergers considered in this work. Their tracks cross the *LISA* sensitivity curve around $f_{\text{min}} \approx$ a few 10^{-4} Hz where the analytical and numerical sensitivity curves match extremely well.

4.3 *LISA* detectability of GW from the high- z Universe

To confidently claim detection, the SNR of an event must be above a critical value. Here, we adopt the typical *LISA* threshold of $\text{SNR} = 7$. Each row in Fig. 8 represents the calculated SNR values for all simulated binaries in a given model as a function of their total intrinsic mass and the redshift. Starting with the fiducial model (*ins1*; top panels), BH mergers become detectable once they reach masses of $\sim 10^4 M_\odot$ at $z \lesssim 13$. As masses grow with time, these systems can reach SNR values as high as ~ 1000 for a total BH mass around $10^5 M_\odot$ below $z \sim 11$. Binaries with $\text{SNR} > 7$ appear in the redshift range $z \simeq 5-13$ and range in total mass between

$M \simeq 10^{3.5-5.6} M_\odot$. Allowing a precise estimation of parameters such as distance, sky localization, and chirp mass, these mass and z ranges will therefore be best probed using GWs. Finally, as the BHs grow above $\sim 10^6 M_\odot$, the SNR decreases as the emitted GW signal shifts to lower frequencies, and above $\sim 10^7 M_\odot$ it goes out of the detectability window. While the results remain quite similar for the *tdf4* model, a delay in the merger time-scales results in a severe reduction in the number of type 2 mergers as shown from the lower rightmost panel of the same figure. Moreover, the detectability window around $10^{4-6} M_\odot$ shifts to slightly lower redshifts. We end by noting that type 2 mergers are rarer than type 1 in both models, with type 3 mergers being the rarest as expected (see Table 2) – this is why type 3 mergers are not plotted here.

We now discuss the yearly high- z event detection rate expected from *LISA* using an $\text{SNR} > 7$. Once the BH merger rate density (per unit comoving volume) of events with $\text{SNR} > 7$ at a given z , $N_{\text{com}}(z)$, is obtained, we convert this into the expected number of mergers per year $d^2N/dzdt$ as (Haehnelt 1994; Arun et al. 2009)

$$\frac{d^2N}{dzdt} = 4\pi c N_{\text{com}}(z) \left(\frac{d_L(z)}{(1+z)} \right)^2 [\text{yr}^{-1}], \quad (31)$$

Table 2. Total number of *LISA* detections expected for an $\text{SNR} > 7$ at $z > 4$ over a 4-yr duration of the mission for the two models that bracket the upper and lower limits of the physical parameter space: *ins1* and *tdf4* for both light and heavy DCBH seeds. We show results for the three different types of BH mergers explained in Section 4.2: (i) *type1 – stellar BH mergers*: mergers of two stellar BH seeds; (ii) *type 2 – mixed mergers*: mergers of a stellar BH seed with a DCBH, and (iii) *type 3 – DCBH–DCBH mergers*: extremely rare, these are mergers of two DCBH seeds. While columns 2–5 show results for all type 1, 2, and 3 mergers, respectively, for an $\text{SNR} > 7$, columns 6–9 show results for the same quantities without imposing any SNR cut.

Model	All ₇	Type 1 ₇	Type 2 ₇	Type 3 ₇	All ₀	Type 1 ₀	Type 2 ₀	Type 3 ₀
<i>ins1</i>	19.8	13	6.8	0.05	300.3	288.3	11.8	0.15
<i>tdf4</i>	12.5	12.1	0.4	0	247.7	247.0	0.62	0.01
<i>ins1</i> (heavy)	23.3	13	10.3	0.04	300.3	288.3	11.8	0.15
<i>tdf4</i> (heavy)	12.5	12.1	0.4	0	247.7	247.0	0.62	0.01

where $d_L(z)$ is the luminosity distance at z . The results of this calculation are shown in Fig. 9. As shown, the fraction of *LISA*-detectable events rises with decreasing redshift from about 1/400 at $z \simeq 13$ to about 1/10 by $z \simeq 8$ to as high as 1/4 by $z \simeq 5$. As expected, most of these events are type 1 mergers. Quantitatively, by $z \simeq 4$, roughly 66 per cent of detectable mergers are type 1 with about 32 per cent being type 2 mergers, with type 3 mergers only contributing 0.3 per cent to the total number. While the qualitative behaviour is quite similar in the *tdf4* case, given the slower BH mass growth, type 1 mergers significantly increase (contributing about 96 per cent to the cumulative event rate by $z \simeq 4$) while the contribution of type 2 mergers falls to roughly 3 per cent. Crucially, we do not find any type 3 mergers above the detection limit in this case.

We find that considering the ‘heavy DCBH seed’ model leads to a slight change in these numbers for the *ins1* case: while the cumulative contribution of type 1 mergers drops slightly to 52 per cent, this is compensated by an increase (to 47 per cent) in the cumulative number of detectable type 2 mergers while the number of type 3 mergers remains unchanged. This heavier seed model, however, has no impact on the results from the *tdf4* model.

The *total* number of detections per model and merger type for the *LISA* mission (over 4 yr) are summarized in Table 2. The model *ins1* with ‘heavy DCBH seeds’ yields the highest total detection number of ~ 23 events comprising ~ 13 type 1 and ~ 10 type 2 mergers. These numbers reduce slightly to about 20 total events comprising 13 type 1 and 7 type 2 mergers using the ‘light DCBH seed’ model. In contrast, only a dozen events (all of type 1) are expected using model *tdf4*; as expected from the discussion above, the DCBH seed mass has no bearing on these results.

We also calculate the event rate in terms of the redshifted merged mass, $M_z = M(1+z)$, such that

$$\frac{d^2 N}{dM_z dt} = 4\pi c N_{\text{com}}(M_z) \left(\frac{d_L(z)}{(1+z)} \right)^2 [\text{yr}^{-1}]. \quad (32)$$

The results of this calculation, presented in Fig. 10, clearly show the *LISA* detectability preference for BH masses ranging between 10^4 and $10^7 M_\odot$ for type 1 and type 2 mergers for both the *ins1* and *tdf4* models. Type 3 mergers, instead, are detectable in the mass range $10^{5-7} M_\odot$ in the ‘light’ DCBH seed model while being undetectable in the *tdf4* model. Moving on to the ‘heavy DCBH seed model’, while the mass range remains unchanged for type 1 mergers, the range for both type 2 and type 3 mergers decreases: while the former range between $10^{5-7} M_\odot$ for both the *ins1* and *tdf4* models, the type 3 range lies in the very narrow range of $10^{5.5-6.5} M_\odot$ for the *ins1* case; as expected, the numbers of mergers of each type in each model are similar to the cumulative numbers quoted above. Practically, however, it would be difficult to distinguish

between these different seeding models purely from the detected mass function, given that all types of merger reside in the same mass range between $10^{4-7} M_\odot$.

Finally, we provide a comparison of our expected event rates with those available in the literature: starting with heavy seeds, all previous studies used DCBH models based on ‘dynamical’ instabilities, of the type advocated by Begelman et al. (2006), Lodato & Natarajan (2006), and Volonteri & Begelman (2010). In this study, we have focused on the currently favoured (at least by the first star community) ‘thermodynamical’ models, which require a high level of LW background for the formation of seeds. As shown in this paper and in Habouzit et al. (2016), this model results in much rarer seeds. We find that Klein et al. (2016) predict 3.9 mergers/year at $z > 4$ in the Q3-d model based on Lodato & Natarajan (2006); note that using the same seeding model as Klein et al. (2016), Bonetti et al. (2018) obtain results consistent with previous literature. Further, Sesana et al. (2007) predict 2.2 mergers/year at $z > 4$ in the model based on Begelman et al. (2006), while the LW-based model explored in this paper yields 0.0025–0.035 mergers/year at $z > 4$ as shown in Table 2. A comparison with Ricarte & Natarajan (2018), who also use a model based on Lodato & Natarajan (2006) and do not include an LW condition, is more difficult because they show only events with $\text{SNR} > 5$. Using this SNR cut, we find that the peak in the rates for heavy seeds is similarly broad and covers a similar redshift range when comparing our results to theirs although they predict a larger number of events: their peak rate is between 0.5 and 5 events/year while our peak rate goes from ~ 0.05 to 0.25 events/year. To summarize, our lower merger rates for heavy seeds, compared to previous works, are what should be expected for a model that predicts extremely rare seeds. This is the effect of adding the condition on the LW background that previous models had not included.

For light (popIII) seeds, Klein et al. (2016) predict 146.3 mergers/year at $z > 4$ (although they extrapolate to 2x this rate in their table 1 and related text) and Sesana et al. (2007) predict 57.7 mergers/year at $z > 4$. Our model predicts between 62.0 and 75.1 mergers/year at $z > 4$ as shown in Table 2. When comparing to Ricarte & Natarajan (2018), again the peak in the rates for light seeds is similarly broad and covers a similar range in redshift but the values of the peak rates are lower in our case. In particular, our type 1 peak rate is 0.75 event/year in the optimistic (*ins1*) and pessimistic (*tdf4*) models, while in Fig. 9 their peak rates lie between ~ 5 and 20 events/year. While our merger rate for light seeds is well within the expectations of the literature, as we made similar assumptions, the results being on the lower side are likely because of the resolution of our merger trees: for instance, Ricarte & Natarajan (2018) have a mass resolution of $\sim 10^6 M_\odot$, while Klein et al. (2016) follow Barausse (2012), who follows Volonteri et al. (2003) (whose trees

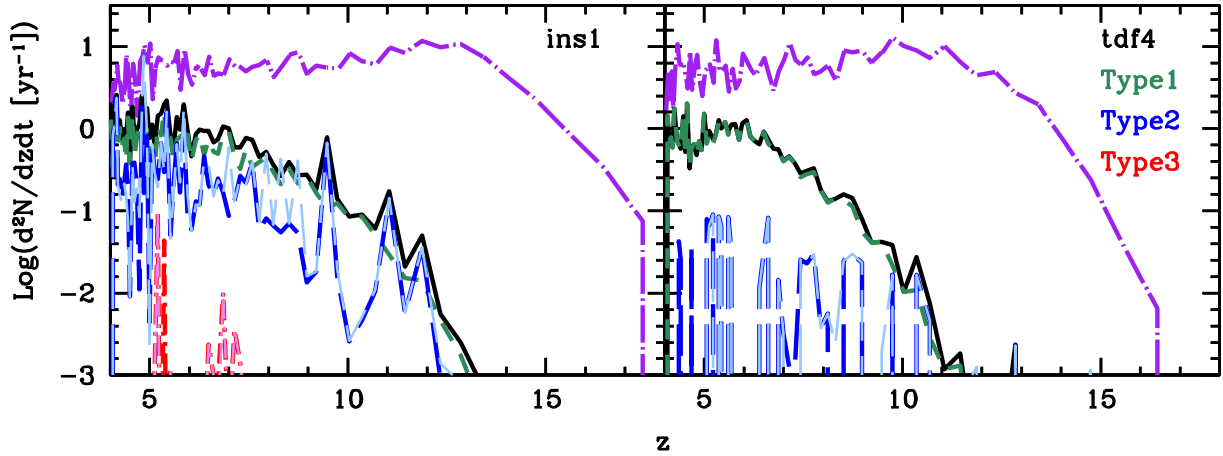


Figure 9. The BH merger event rate (per year) expected as a function of redshift for two models that bracket the physical range probed: *left-hand panel*: ins1 and *right-hand panel*: tdf4. In each panel, the dot-dashed purple line shows the results for all mergers (without any cut in SNR) while the solid black line shows the results for all mergers using a value of $\text{SNR} > 7$. The latter is deconstructed into the contribution from (SNR > 7) type1 (green dashed line), type2 ‘light DCBH’ seed (dark blue dashed line), and type3 ‘light DCBH’ seed (red dashed line) mergers. Further, the long-dashed light blue line and dot-dashed pink line show results for mergers with SNR > 7 using a heavier DCBH seed mass of $10^{4-5} M_{\odot}$ for type 2 and type 3 mergers, respectively. These results are in general agreement with those used for *LISA* calculations (e.g. Fig. 3, Klein et al. 2016).

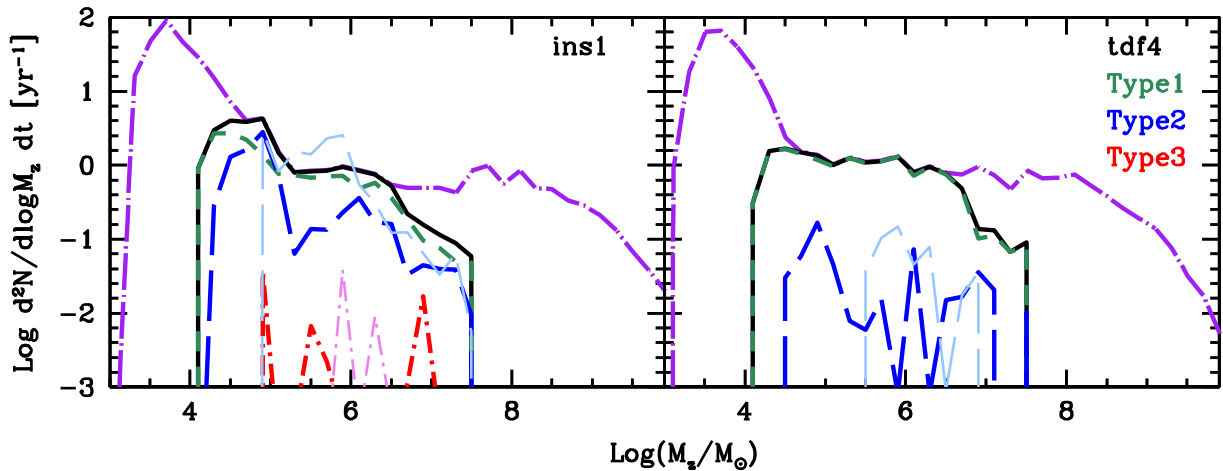


Figure 10. The BH merger event rate (per year) as a function of the redshifted BH mass [$M_z = M_{\text{bh}}(1 + z)$]. The lines show the same models as noted in Fig. 9.

are used for Sesana et al. 2007) in having a resolution dependent on the halo mass at $z = 0$, reaching $10^5 M_{\odot}$ for haloes with a mass $< 4 \cdot 10^{12} M_{\odot}$ at $z = 0$ and up to $10^7 M_{\odot}$ for haloes with a mass $10^{15} M_{\odot}$ at $z = 0$.

5 CONCLUSIONS AND DISCUSSION

In this work, we have included the impact of BH seeding, growth, and feedback into our semi-analytic model, *Delphi*. Our model now jointly tracks the build-up of the dark matter halo, gas, stellar, and BH masses of high- z ($z \gtrsim 5$) galaxies. We remind the reader that our star formation efficiency is the minimum between the star formation rate that equals the halo binding energy and a saturation efficiency. In the same flavour, the BH accretion at any time-step is the minimum between the BH accreting a certain fraction of the gas mass leftover after star formation, up to a fraction of the Eddington limit: while high-mass haloes can accrete at the Eddington limit,

low-mass haloes follow a lower efficiency track. We explore a number of physical scenarios using this model that include (i) two types of BH seeds (stellar and those from DCBH); (ii) the impact of reionization; and (iii) the impact of instantaneous versus delayed galaxy mergers on the baryonic growth.

We show that, using a minimal set of mass- and z -independent free parameters, our model reproduces all available data sets for high- z galaxies and BH including the evolving (galaxy and AGN) UVLF, the SMD, and the BHMF. Crucially, our model naturally yields a BH mass–stellar mass relation that is tightly coupled for high-stellar mass ($M_* \gtrsim 10^{9.5} M_{\odot}$) haloes; lower mass haloes, on the other hand, show a stunted BH growth. Interestingly, while both reionization feedback and delayed mergers have no impact on the UVLF, the SMD is more affected by the reionization feedback as compared to the delayed mergers.

We then use this model, benchmarked against all available high- z data, to predict the merger event rate expected for the *LISA* mission.

We find that *LISA*-detectable binaries (with $\text{SNR} > 7$) appear in the redshift range $z \simeq 5\text{--}13$ and range in total mass between $M \simeq 10^{3.5\text{--}5} M_{\odot}$. While type 1 mergers (of two stellar BHs) dominate in all the scenarios studied, type 2 mergers (merger of a stellar BH and a DCBH) can contribute as much as 32 percent to the cumulative event rate by $z \sim 4$ in the fiducial (*ins1*) model. However, including the impact of reionization feedback and delayed mergers (*tdf4* model) results in a lower BH growth with type 2 mergers contributing only 3 percent to the cumulative event rates. Using heavier DCBH seeds results in a larger number of type 2 mergers becoming detectable with *LISA* while leaving the results effectively unchanged for the *tdf4* model.

Quantitatively, the model *ins1* with ‘heavy DCBH seeds’ yields the highest total detection number of ~ 23 events comprising ~ 13 type 1 and ~ 10 type 2 mergers. These numbers reduce slightly to about 20 total events comprising 13 type 1 and 7 type 2 mergers using the ‘light DCBH seed’ model. In contrast, only a dozen events (all of type 1) are expected using model *tdf4* and the DCBH seed mass has no bearing on these results.

We end with a few caveats. First, given that we do not consider (the realistic case of) recoil and BH ejection from the host haloes, all BHs remain bound to haloes. Secondly, the enhancement of the LW seen by any haloes only depends on its bias at that redshift. This effectively means that we ignore the impact of the local environment on the LW intensity seen by any halo, and this may lead to an increase in seed formation and mergers in more biased regions. Thirdly, we have not included BH seeds from stellar dynamical channels that have a milder metallicity dependence and should have a number density intermediate between SBHs and DCBHs (e.g. Devecchi et al. 2012; Lupi et al. 2014); DCBH models that are metallicity-independent can also provide an additional channel, increasing the BH merger rate over cosmic time (Volonteri & Begelman 2010; Bonoli, Mayer & Callegari 2014). Finally, we have used a very crude mode for reionization feedback that ignores the patchiness of reionization – in our model, haloes either remain unaffected by the UVB or haloes below a certain chosen virial velocity have all of their gas mass completely photoevaporated. We aim to address each of these intricacies in detail in future works.

ACKNOWLEDGEMENTS

PD acknowledges support from the European Research Council’s starting grant ERC StG-717001 (‘DELPHI’). PD and OP acknowledge support from the European Commission’s and University of Groningen’s CO-FUND Rosalind Franklin programme. MV acknowledges funding from the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007–2013 Grant Agreement no. 614199, project ‘BLACK’). Finally, PD thanks A. Mazumdar for his scientific inputs, which have greatly added to the paper.

REFERENCES

Abel T., Bryan G. L., Norman M. L., 2002, *Science*, 295, 93
 Agarwal B. et al., 2018, preprint (arXiv:1807.08499)
 Agarwal B., Dalla Vecchia C., Johnson J. L., Khochfar S., Paardekooper J.-P., 2014, *MNRAS*, 443, 648
 Amaro-Seoane P. et al., 2012, *Class. Quantum Gravity*, 29, 124016
 Amaro-Seoane P. et al., 2017, preprint (arXiv:1702.00786)
 Anglés-Alcázar D., Faucher-Giguère C.-A., Quataert E., Hopkins P. F., Feldmann R., Torrey P., Wetzel A., Kereš D., 2017, *MNRAS*, 472, L109
 Arun K. G. et al., 2009, *Class. Quantum Gravity*, 26, 094027

Atek H. et al., 2015, *ApJ*, 814, 69
 Barack L. et al., 2018, preprint (arXiv:1806.05195)
 Barausse E., 2012, *MNRAS*, 423, 2533
 Begelman M. C., Volonteri M., Rees M. J., 2006, *MNRAS*, 370, 289
 Begelman M. C., Rossi E. M., Armitage P. J., 2008, *MNRAS*, 387, 1649
 Bonetti M., Sesana A., Haardt F., Barausse E., Colpi M., 2018, preprint (arXiv:1812.01011)
 Bonoli S., Mayer L., Callegari S., 2014, *MNRAS*, 437, 1576
 Bouwens R. J. et al., 2010, *ApJ*, 725, 1587
 Bouwens R. J. et al., 2015, *ApJ*, 803, 34
 Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2007, *ApJ*, 670, 928
 Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, *MNRAS*, 465, 32
 Bowler R. A. A. et al., 2014, *MNRAS*, 440, 2810
 Bradley L. D. et al., 2012, *ApJ*, 760, 108
 Bromm V., Loeb A., 2003, *ApJ*, 596, 34
 Castellano M. et al., 2010, *A&A*, 524, A28
 Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, 319, 168
 Colpi M., 2018, preprint (arXiv:1807.06967)
 Dayal P., Ferrara A., 2018, *Phys. Rep.*, 780, 1
 Dayal P., Dunlop J. S., Maio U., Ciardi B., 2013, *MNRAS*, 434, 1486
 Dayal P., Ferrara A., Dunlop J. S., Pacucci F., 2014, *MNRAS*, 445, 2545
 Dayal P., Mesinger A., Pacucci F., 2015, *ApJ*, 806, 67
 Dayal P., Choudhury T. R., Bromm V., Pacucci F., 2017a, *ApJ*, 836, 16
 Dayal P., Choudhury T. R., Pacucci F., Bromm V., 2017b, *MNRAS*, 472, 4414
 Decarli R. et al., 2018, *ApJ*, 854, 97
 Devecchi B., Volonteri M., 2009, *ApJ*, 694, 302
 Devecchi B., Volonteri M., Rossi E. M., Colpi M., Portegies Zwart S., 2012, *MNRAS*, 421, 1465
 Dijkstra M., Ferrara A., Mesinger A., 2014, *MNRAS*, 442, 2036
 Dotan C., Rossi E. M., Shaviv N. J., 2011, *MNRAS*, 417, 3035
 Dubois Y., Volonteri M., Silk J., Devriendt J., Slyz A., Teyssier R., 2015, *MNRAS*, 452, 1502
 Duncan K. et al., 2014, *MNRAS*, 444, 2960
 Ferrara A., Salvadori S., Yue B., Schleicher D., 2014, *MNRAS*, 443, 2410
 Fiacconi D., Rossi E. M., 2016, *MNRAS*, 455, 2
 Fiacconi D., Rossi E. M., 2017, *MNRAS*, 464, 2259
 González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, *ApJ*, 735, L34
 Grazian A. et al., 2015, *A&A*, 575, A96
 Habouzit M., Volonteri M., Dubois Y., 2017, *MNRAS*, 468, 3935
 Habouzit M., Volonteri M., Latif M., Dubois Y., Peirani S., 2016, *MNRAS*, 463, 529
 Haehnelt M. G., 1994, *MNRAS*, 269, 199
 Hartwig T., 2018, preprint (arXiv:1807.06155)
 Hartwig T., Agarwal B., Regan J. A., 2018, *MNRAS*, 479, L23
 Jeon M., Pawlik A. H., Bromm V., Milosavljević M., 2014, *MNRAS*, 440, 3778
 Jiang L. et al., 2016, *ApJ*, 833, 222
 Johnson J. L., Whalen D. J., Fryer C. L., Li H., 2012, *ApJ*, 750, 66
 Kashikawa N. et al., 2015, *ApJ*, 798, 28
 Klein A. et al., 2016, *Phys. Rev. D*, 93, 024003
 Kulkarni G., Worsack G., Hennawi J. F., 2018, preprint (arXiv:1807.09774)
 Labbé I. et al., 2013, *ApJ*, 777, L19
 Lacey C., Cole S., 1993, *MNRAS*, 262, 627
 Latif M. A., 2018, preprint (arXiv:1807.06337)
 Latif M. A., Ferrara A., 2016, *Publ. Astron. Soc. Aust.*, 33, e051
 Latif M. A., Schleicher D. R. G., Schmidt W., Niemeyer J., 2013, *MNRAS*, 433, 1607
 Lauer T. R., Tremaine S., Richstone D., Faber S. M., 2007, *ApJ*, 670, 249
 Leitherer C. et al., 1999, *ApJS*, 123, 3
 Livermore R. C., Finkelstein S. L., Lotz J. M., 2017, *ApJ*, 835, 113
 Lodato G., Natarajan P., 2006, *MNRAS*, 371, 1813
 Loeb A., Rasio F. A., 1994, *ApJ*, 432, 52
 Lupi A., Colpi M., Devecchi B., Galanti G., Volonteri M., 2014, *MNRAS*, 442, 3616
 Madau P., Rees M. J., 2001, *ApJ*, 551, L27

- Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, *MNRAS*, 351, 169
- McGreer I. D. et al., 2013, *ApJ*, 768, 105
- McLure R. J. et al., 2013, *MNRAS*, 432, 2696
- McLure R. J., Cirasuolo M., Dunlop J. S., Foucaud S., Almaini O., 2009, *MNRAS*, 395, 2196
- McLure R. J., Dunlop J. S., Cirasuolo M., Koekemoer A. M., Sabbi E., Stark D. P., Targett T. A., Ellis R. S., 2010, *MNRAS*, 403, 960
- Oesch P. A. et al., 2013, *ApJ*, 773, 75
- Oesch P. A. et al., 2014, *ApJ*, 786, 108
- Oke J. B., Gunn J. E., 1983, *ApJ*, 266, 713
- Ono Y. et al., 2018, *PASJ*, 70, S10
- Parsa S., Dunlop J. S., McLure R. J., 2018, *MNRAS*, 474, 2904
- Pezzulli E., Valiante R., Schneider R., 2016, *MNRAS*, 458, 3047
- Planck Collaboration, 2016, *A&A*, 594, 13
- Regan J. A., Haehnelt M. G., 2009, *MNRAS*, 396, 343
- Ricarte A., Natarajan P., 2018, *MNRAS*, 481, 3278
- Salviander S., Shields G. A., Gebhardt K., Bonning E. W., 2007, *ApJ*, 662, 131
- Santamaría L. et al., 2010, *Phys. Rev. D*, 82, 064016
- Sesana A., Volonteri M., Haardt F., 2007, *MNRAS*, 377, 1711
- Sesana A., Gair J., Berti E., Volonteri M., 2011, *Phys. Rev. D*, 83, 044036
- Shang C., Bryan G. L., Haiman Z., 2010, *MNRAS*, 402, 1249
- Shao Y. et al., 2017, *ApJ*, 845, 138
- Smith B. D., Regan J. A., Downes T. P., Norman M. L., O’Shea B. W., Wise J. H., 2018, *MNRAS*, 480, 3762
- Song M. et al., 2016, *ApJ*, 825, 5
- Stark D. P., Schenker M. A., Ellis R., Robertson B., McLure R., Dunlop J., 2013, *ApJ*, 763, 129
- Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, *ApJ*, 519, 1
- Sugimura K., Omukai K., Inoue A. K., 2014, *MNRAS*, 445, 544
- Tanaka T., Haiman Z., 2009, *ApJ*, 696, 1798
- Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, *ApJ*, 786, 104
- Valiante R., Schneider R., Volonteri M., Omukai K., 2016, *MNRAS*, 457, 3356
- Venemans B. P., Walter F., Zschaechner L., Decarli R., De Rosa G., Findlay J. R., McMahon R. G., Sutherland W. J., 2016, *ApJ*, 816, 37
- Vito F. et al., 2018, *MNRAS*, 473, 2378
- Volonteri M., Begelman M. C., 2010, *MNRAS*, 409, 1022
- Volonteri M., Reines A. E., 2016, *ApJ*, 820, L6
- Volonteri M., Stark D. P., 2011, *MNRAS*, 417, 2085
- Volonteri M., Haardt F., Madau P., 2003, *ApJ*, 582, 559
- Willott C. J. et al., 2010, *AJ*, 140, 546

This paper has been typeset from a \LaTeX file prepared by the author.